



IDICSO

Instituto de Investigación en Ciencias Sociales
Facultad de Ciencias Sociales
Universidad del Salvador

ÁREA EMPLEO Y POBLACIÓN

El análisis de las series cronológicas

por **Horacio Chitarroni***

Buenos Aires, DIC/2002

* **CHITARRONI, Horacio.** Lic. en Sociología, Universidad Nacional de Buenos Aires (UBA). Docente, Facultad de Ciencias Sociales, Universidad del Salvador (USAL). Docente de la Maestría en Ciencias Sociales del Trabajo, Facultad de Ciencias Sociales, UBA. Investigador Principal, Área Empleo y Población, IDICSO, USAL. Consultor del Consejo Nacional de Coordinación de Políticas Sociales, SIEMPRO (Sistema de Evaluación, Seguimiento y Monitoreo de Programas Sociales).

BREVE HISTORIA DEL IDICSO. Los orígenes del IDICSO se remontan a 1970, cuando se crea el "Proyecto de Estudio sobre la Ciencia Latinoamericana (ECLA)" que, por una Resolución Rectoral (21/MAY/1973), adquiere rango de Instituto en 1973. Desde ese entonces y hasta 1981, se desarrolla una ininterrumpida labor de investigación, capacitación y asistencia técnica en la que se destacan: estudios acerca de la relación entre el sistema científico-tecnológico y el sector productivo, estudios acerca de la productividad de las organizaciones científicas y evaluación de proyectos, estudios sobre política y planificación científico tecnológica y estudios sobre innovación y cambio tecnológico en empresas. Las actividades de investigación en esta etapa se reflejan en la nómina de publicaciones de la "Serie ECLA" (SECLA). Este instituto pasa a depender orgánica y funcionalmente de la Facultad de Ciencias Sociales a partir del 19 de Noviembre de 1981, cambiando su denominación por la de Instituto de Investigación en Ciencias Sociales (IDICSO) el 28 de Junio de 1982.

Los fundamentos de la creación del IDICSO se encuentran en la necesidad de:

- ❖ Desarrollar la investigación pura y aplicada en Ciencias Sociales.
- ❖ Contribuir a través de la investigación científica al conocimiento y solución de los problemas de la sociedad contemporánea.
- ❖ Favorecer la labor interdisciplinaria en el campo de las Ciencias Sociales.
- ❖ Vincular efectivamente la actividad docente con la de investigación en el ámbito de la facultad, promoviendo la formación como investigadores, tanto de docentes como de alumnos.
- ❖ Realizar actividades de investigación aplicada y de asistencia técnica que permitan establecer lazos con la comunidad.

A partir de 1983 y hasta 1987 se desarrollan actividades de investigación y extensión en relación con la temática de la integración latinoamericana como consecuencia de la incorporación al IDICSO del Instituto de Hispanoamérica perteneciente a la Universidad del Salvador. Asimismo, en este período el IDICSO desarrolló una intensa labor en la docencia de post-grado, particularmente en los Doctorados en Ciencia Política y en Relaciones Internacionales que se dictan en la Facultad de Ciencias Sociales. Desde 1989 y hasta el año 2001, se suman investigaciones en otras áreas de la Sociología y la Ciencia Política que se reflejan en las series "Papeles" (SPI) e "Investigaciones" (SII) del IDICSO. Asimismo, se llevan a cabo actividades de asesoramiento y consultoría con organismos públicos y privados. Sumándose a partir del año 2003 la "Serie Documentos de Trabajo" (SDTI).

La investigación constituye un componente indispensable de la actividad universitaria. En la presente etapa, el IDICSO se propone no sólo continuar con las líneas de investigación existentes sino también incorporar otras con el propósito de dar cuenta de la diversidad disciplinaria, teórica y metodológica de la Facultad de Ciencias Sociales. En este sentido, las líneas de investigación del IDICSO constituyen ámbitos de articulación de la docencia y la investigación así como de realización de tesis de grado y post-grado. En su carácter de Instituto de Investigación de la Facultad de Ciencias Sociales de la Universidad del Salvador, el IDICSO atiende asimismo demandas institucionales de organismos públicos, privados y del tercer sector en proyectos de investigación y asistencia técnica.

IDICSO

Departamento de Comunicación

Email: idicso@yahoo.com.ar

Web Site: <http://www.salvador.edu.ar/csoc/idicso>

Introducción

Una serie cronológica es, básicamente, una variable Y cuyas observaciones ($Y_1, Y_2, y_3...Y_n$) han sido tomadas en distintos momentos del tiempo: $t_1, t_2, t_3...t_n$. Se dice que Y es una función de t , lo que sugiere que las variaciones de Y están causadas por las variaciones de t . Sin embargo, aunque Y varíe así como transcurre el tiempo, es obvio que sus cambios de valores no ocurren meramente por ello, sino que son, seguramente, causados por otros fenómenos –no conocidos– que también experimentan variaciones a medida que pasa el tiempo.

Vale decir, en lugar de haber medido el desempleo para las diferentes provincias argentinas, medimos el desempleo de Córdoba a lo largo de varias ondas de la EPH. Es, pues, como si tuviéramos varias Córdobas: digamos, desde la de mayo de 1998 a la de mayo de 2002. En ese lapso, el desempleo aumentó: no meramente por el paso de los años, sino influenciado por ciertos acontecimientos que ocurrían en forma paralela en la economía y el mercado de trabajo.

El estudio de las variaciones de una variable en el tiempo responde básicamente a tres propósitos (que, en rigor, son los propósitos habituales de la actividad científica):

- ❖ podemos desear describir el comportamiento de esa variable: ¿aumentó o descendió el desempleo en el lapso considerado?, ¿lo hizo mucho o poco, por ejemplo, en comparación con otros períodos previos, o bien con otras ciudades durante el mismo lapso?
- ❖ podemos querer descubrir algunas eventuales causas de tales variaciones (los factores “encubiertos” tras el paso del tiempo): ¿cómo se comportó la tasa de actividad en ese período?, ¿cuál fue la evolución de ciertas variables como el PIB, la inversión o el consumo privado en el mismo lapso?
- ❖ o bien –más ambiciosamente– podemos pretender predecir el comportamiento futuro de la variable: de mantenerse relativamente constante el escenario actual, ¿cuál sería la evolución presumible del desempleo en los próximos meses o en el año entrante?

El último propósito se basa en un supuesto epistemológico algo burdo y pragmático pero útil: aquel que dice que así como fueron las cosas hasta ahora, seguirán siendo en el futuro... (esto último se basa en la asunción de que los fenómenos no ocurren por mero azar, sino guiados por unas ciertas leyes inmanentes).

Los componentes de las series temporales

A los efectos de su análisis (en especial cuando nos animan propósitos explicativos y predictivos) una serie temporal puede ser analíticamente descompuesta en cuatro componentes:

$$Y = T * C * E * I$$

- ❖ Una tendencia secular o de largo plazo (T)
- ❖ Unas variaciones cíclicas (C)
- ❖ Unas variaciones estacionales (E)
- ❖ Unas variaciones irregulares (I)

Es fácil ejemplificar cada uno de estos componentes que, en general, resultarán familiares:

- ❖ Hay variables que tienen una tendencia o comportamiento de largo plazo: por ejemplo, el producto bruto mundial tiende a crecer en el largo plazo, por mera agregación de factores: cada vez hay más agentes que producen y consumen. En cambio, algunas variables como las tasas de mortalidad infantil tienden a decrecer, porque mejoran a largo plazo las condiciones sanitarias, el acceso al agua potable, la vacunación, etc. (hasta por el mero efecto de la urbanización)
- ❖ Otras variables (tengan o no un comportamiento secular o de largo plazo) reconocen movimientos cíclicos: el ejemplo más corriente es el de los ciclos económicos de auge y recesión. A lo largo de la tendencia a crecer del producto bruto mundial, hay ondulaciones: ciclos ascendentes y descendentes de variable duración e intensidad
- ❖ Asimismo, dentro de los ciclos hay, a veces, unos movimientos de más corto plazo, repetidos y regulares, que se producen habitualmente para las mismas épocas: se trata de las estacionalidades. Por ejemplo, el gasto en turismo es estacional: independientemente del comportamiento del ciclo (es decir, tanto en épocas de recesión como de prosperidad) aumenta en verano y en vacaciones de invierno, y desciende en “temporada baja” (abril/junio y agosto/noviembre). Del mismo modo que la venta de gaseosas o de helados aumenta en verano y ello sucede todos los años, indefectiblemente.
- ❖ Finalmente, hay cambios en el comportamiento de ciertas variables que son producidos por ciertos eventos inusuales. Por ejemplo, un fenómeno meteorológico puede arruinar una cosecha y alterar la evolución de los precios de cierto producto (al reducir su oferta), el atentado a las Torres Gemelas produjo una alteración en ciertas variables económicas o el resultado de una elección puede modificar el volumen de los flujos de inversión hacia un país. Este tipo de variaciones, justamente en razón de su irregularidad, son poco predecibles.

Por cierto que todos estos componentes no están necesariamente presentes: hay variables que carecen de estacionalidad o de ciclo, o que no están gobernadas por una tendencia de largo plazo (al menos en forma clara y perceptible).

Existen procedimientos que permiten “capturar” estos componentes, aislarlos y despojar a las variables de ellos, para ver cómo se comportarían si alguno de

ellos no operara, o bien si sólo operara uno en particular. La descomposición de una serie en estos componentes puede ser útil tanto para tratar de predecir su comportamiento como para explicar las causas de sus variaciones.

Por ejemplo, si en la ecuación $Y = T * C * E * I$ pudiéramos eliminar la estacionalidad y la tendencia, las variaciones de Y que subsistieran serían imputables a los dos componentes restantes. Un par de ejemplos permitirán aclarar esto: supongamos que se analiza la evolución de ciertos indicadores (podría ser justamente la tasa de desempleo) para determinar si se está revirtiendo el ciclo económico. Pues bien: muchos estudiosos afirman que la tasa (que se mide en los meses de mayo y octubre de cada año) está afectada por cierta estacionalidad. Podría, también, tener una tendencia –por ejemplo, creciente– de largo plazo. Todo esto afectaría sus variaciones y dificultaría apreciar el comportamiento en relación con el ciclo: si pudiéramos eliminar los efectos de la tendencia y la estacionalidad –mejor aún, si también pudiéramos, de algún modo, controlar las variaciones irregulares– quedaría al desnudo el comportamiento cíclico.

Un estudio de Frenkel y Sánchez Rosada acerca del desempleo industrial durante la convertibilidad tenía como propósito determinar la influencia de la reducción del arancel de importación en ciertas ramas de la industria. En este caso, una reducción arancelaria es una medida de política económica: un evento irregular que ocurre en cierto momento. Para constatar sus efectos haría falta desestacionalizar la tasa, despojarla de la tendencia y –si fuera posible– de las variaciones cíclicas.

En cambio, cuando se trata de predecir el comportamiento futuro, a largo plazo, lo más atinado es tratar de valerse de la tendencia, para hacer proyecciones: esto es difícil en el caso de variables como el desempleo, donde es mayor la influencia del resto de los componentes. Pero, en cambio, es muy factible cuando se trata de variables muy estructurales, como es el caso de las demográficas: aún antes del último censo el INDEC disponía de proyecciones de población razonablemente ajustadas, por sexo y edad, que llegaban al 2025. Esto es posible, porque los comportamientos reproductivos y las tasas de mortalidad de una población no sufren alteraciones bruscas (salvo cuando ocurren sucesos irregulares tales como una guerra prolongada, un desastre natural de enorme magnitud o un impacto migratorio muy intenso).

Algunos procedimientos descriptivos

Existen algunos procedimientos muy sencillos que facilitan el análisis descriptivo de las series cronológicas. Uno de ellos es el llamado **índice base 100**. Si disponemos de información acerca de la evolución de los ingresos laborales a lo largo de un cierto período, resulta más fácil apreciar esa evolución si expresamos todos los datos en función de un año cualquiera, tomado como base = 100. Este año puede ser el año inicial de la serie, el año final, o bien un año intermedio que tengamos buenas razones para seleccionar. Para hacerlo, basta con dividir cada registro por el correspondiente al año base y multiplicarlo

luego por 100. Este procedimiento no tiene otro objeto que facilitar una rápida comparación¹.

Total de aglomerados urbanos: evolución de los ingresos laborales (a pesos corrientes) 1998/2002

Período	Mayo1998	Mayo 1999	Mayo 2000	Mayo 2001	Mayo 2002
A pesos corrientes	614	583	554	540	449
Índice base 1998 = 100	100,0	95,0	90,2	87,9	73,1

Fuente: EPH-INDEC (ondas de mayo de cada año)

Fácilmente se advierte que el ingreso laboral era, en mayo de 2002, equivalente a 73% del correspondiente a igual mes de 1998: vale decir, experimentó – en términos nominales – una caída de 27%.

Un segundo procedimiento consiste en calcular las tasas de variación interperíodos, lo que se hace mediante una fórmula muy simple:

$$(VF - VI) / VI * 100$$

Donde :

VF: valor del período t_1

VI: valor del período t_0

Está claro que si el valor final es menor que el inicial, la tasa resulta negativa.

Total de aglomerados urbanos: variación del ingreso laboral nominal. 1998/2002

Período	Mayo1998	Mayo 1999	Mayo 2000	Mayo 2001	Mayo 2002
A pesos corrientes	614	583	554	540	449
Variación (%)		-5,0	-5,0	-2,5	-16,9

Fuente: EPH-INDEC (ondas de mayo de cada año)

El ingreso nominal cayó 5% en los dos primeros períodos, 2,5% en el tercero y casi 17% en el último. Si calculáramos la tasa entre los extremos, comprobaríamos la caída de 27%:

¹ Además, esto supone comparar ingresos nominales, sin considerar la evolución de su poder adquisitivo: esto último requeriría deflactar los ingresos mediante un índice que reflejara la variación de los precios de bienes y servicios, para comparar ingresos reales, a pesos constantes de una cierta fecha. Existe un variado conjunto de índices, como los de Laspeyres, Paasche, Fisher, etc.

$$(449 - 614) / 614 * 100 = -26,9$$

Si las tasas de variación fueran muy constantes (por ejemplo, todas próximas a 5%) sería razonable calcular un promedio de ellas y utilizarlo para proyectar el ingreso para períodos posteriores, simplemente mediante la aplicación repetida de esa tasa promedio: podríamos suponer que el ingreso en mayo de 2003 sería equivalente a 95% del de mayo de 2002 y que volvería a reducirse otro 5% en el 2004. Cosas parecidas a esta (aunque bastante más complejas) hacen los demógrafos para proyectar la población.

Cómo estimar la tendencia

La estimación de la tendencia (T en la ecuación ya vista) de una variable Y puede hacerse de más de un modo. Pero lo más habitual es utilizar una función de regresión (por ejemplo la ya conocida recta de mínimos cuadrados), donde la variable X es el tiempo. Si el ajuste a la recta es bueno (si el coeficiente r de Pearson es razonablemente alto), entonces la recta permitirá estimar valores de Y para t_n , basándose en la ecuación:

$$Y_n = a + b * t_n$$

Es decir, para cada valor observado de Y (Y_o) tendríamos una estimación de Y (Y_e). Si la recta de regresión representa la tendencia (T), las diferencias entre los valores observados y estimados de Y obedecerían a las variaciones de Y provocadas por los otros componentes (C, E, I). Hacemos el supuesto de que si no hubiera ciclos ni estacionalidades ni variaciones irregulares, los valores de Y serían exactamente los que predice la recta².

Una vez obtenidos los valores estimados de Y (Y_e), sería posible despojar a los valores observados de Y (Y_o) de la tendencia, para obtener los que esperaríamos hallar si esta tendencia no operara. Ello no es difícil y puede hacerse en dos sencillos pasos:

1. expresar cada uno de estos valores (Y_e) como base = 100 (dividiéndolos por el valor inicial y multiplicándolos por 100³)
2. dividir cada uno de los valores observados (Y_o) por el correspondiente valor estimado (Y_e), ya convertido en base 100 y multiplicarlo por 100: el resultado será una serie de nuevos valores de Y (que diferirán de los originales) y serían los Y sin tendencia

Veamos un ejemplo:

² Por supuesto, en lugar de una función recta podría emplearse otra que ajustara mejor: logarítmica, exponencial (cuadrática o cúbica), etc.

³ Esta operación no es más que un simple cambio de escala.

Tasas de desempleo 1996/2001

Años	1996	1997	1998	1999	2000	2001
Valores observados (1)	17,1	16,1	15,8	17,4	19,2	21,3
Valores estimados por regresión (2)	17,8	18	18,4	19,3	20,3	22,1
Valores estimados (base valor inicial = 100) (3)	100,0	101,1	103,4	108,4	114,0	124,2
Valores sin tendencia (1/3*100)	17,1	15,9	15,3	16,0	16,8	17,2

Fuente: datos ficticios

Cómo eliminar la estacionalidad

También puede despojarse a una serie de las variaciones provocadas por la estacionalidad. Para ello debe construirse un índice estacional, que capta – justamente – esas variaciones. Veremos, también, un ejemplo simple, basado en las mediciones de mayo y octubre de las tasas de desempleo.

Serie original (tasas de desempleo – total de aglomerados urbanos)						
Años	1996	1997	1998	1999	2000	2001
Mayo	17,1	16,1	13,2	14,5	15,4	16,4
Octubre	17,3	13,7	12,4	13,8	14,7	18,3
Promedio	17,2	14,9	12,8	14,2	15,1	17,4

Fuente: EPH-INDEC (varias ondas)

El primer cuadro nos muestra la serie original de las tasas de desempleo, medidas en los meses de mayo y octubre de cada uno de los años considerados. El primer paso consiste en obtener un promedio simple, para cada año, de las tasas de mayo y octubre.

Índice base promedio = 100							
Años	1996	1997	1998	1999	2000	2001	Índice estacional
Mayo	99,4	108,1	103,1	102,5	102,3	94,5	101,7
Octubre	100,6	91,9	96,9	97,5	97,7	105,5	98,3

En el segundo cuadro, cada tasa se ha expresado en base al promedio del año correspondiente = 100. Para ello, por ejemplo, se divide la tasa de mayo de 1996

por el promedio de 1996 y se multiplica por cien: $17,1 / 17,2 \cdot 100 = 99,4$. Hecho esto, ya puede advertirse a simple vista cierta estacionalidad: los mayos suelen ser superiores a 100, en tanto que los octubres tienden a situarse por debajo de 100.

El paso inmediato consiste en obtener sendos promedios (por fila) de las tasas de mayo y de octubre así expresadas. Estos promedios, que son los índices estacionales, capturan la tendencia alcista de los mayor, así como la tendencia a la baja de los octubres (que en eso, precisamente, consiste la estacionalidad): obsérvese que el promedio de mayo es superior a 100, en tanto que el de octubre es levemente inferior. Si no hubiera estacionalidad, los promedios correspondientes a ambos meses no debieran diferir.

Serie desestacionalizada						
Años	1996	1997	1998	1999	2000	2001
Mayo	16,8	15,8	13,0	14,3	15,1	16,1
Octubre	17,6	13,9	12,6	14,0	14,9	18,6

Finalmente, el tercer cuadro nos muestra las tasas desestacionalizadas: para obtenerlas, basta con dividir cada tasa por el índice estacional que corresponde a su mes y luego multiplicar por 100. Así, para la tasa de mayo de 1996:

$$17,1 / 101,7 \cdot 100 = 16,8.$$

La aplicación del índice estacional opera reduciendo los valores de los meses de mayo (que tienen estacionalidad alta: dividimos por una cifra superior a 100) y elevando los valores de los meses de octubre (que tienen baja estacionalidad: dividimos por una cifra inferior a 100). De este modo, habremos despojado a la serie de sus variaciones estacionales.

Qué hacer con las variaciones irregulares

Por último, podemos ver un sencillo procedimiento que resulta de utilidad para suavizar las variaciones irregulares de una serie. Se trata de el cálculo de promedios móviles.

Periodo	Original	Prom. móviles
May-96	17,1	16,8
Oct-96	17,3	15,7
May-97	16,1	14,3
Oct-97	13,7	13,1
May-98	13,2	13,4

Oct-98	12,4	13,6
May-99	14,5	14,6
Oct-99	13,8	14,6
May-00	15,4	15,5
Oct-00	14,7	16,5
May-01	16,4	
Oct-01	18,3	

El cuadro muestra, nuevamente, las tasas de desempleo medidas para los meses de mayo y octubre. En este caso, se han calculado promedios móviles de orden 3, consistentes en promediar los tres primeros valores (mayo'96, octubre'97 y mayo'97), luego correrse un lugar y promediar nuevamente tres valores (octubre'96, mayo'97 y octubre'97), seguidamente correrse nuevamente un lugar y obtener el promedio: mayo'97, octubre'97, mayo'98, etc. El último promedio móvil que podemos obtener será octubre'00, mayo'01 y octubre'01. Por cierto, al proceder así, las doce observaciones originales habrán quedado reducidas a diez, pero se trata de una serie "suavizada", donde hemos eliminado las variaciones bruscas provocadas por sucesos irregulares.

Los movimientos cíclicos

Hemos dejado, deliberadamente, la consideración de los movimientos cíclicos para el final. La razón estriba en que, generalmente, no hay un modo eficaz de capturarlos o predecir su comportamiento. Si una serie presentara ciclos relativamente regulares, de similar duración y profundidad, entonces ellos podrían ser capturados mediante algún índice construido con una metodología semejante a la empleada para las estacionalidades. Pero toda vez que no sea así –es decir, si los ciclos son variables en duración e intensidad– esto ya no será posible. En consecuencia, la alternativa que resta para el estudio de los ciclos (siempre que se disponga de una serie suficientemente larga) será despojarla de todos los otros componentes (tendencia, estacionalidad, variaciones irregulares): la variabilidad que quede sería atribuible al ciclo. Si graficáramos entonces la serie –y si realmente existiera un ciclo– podríamos examinarlo y obtener conclusiones respecto de sus características.

¿En qué orden?

Ya se ha dicho que no todas las series presentan los mismos componentes. Por lo demás, eliminar unos u otros, así como el orden en que se lo haga, depende de los propósitos últimos del análisis. Si, como se lo expresó más arriba, se tratara de examinar el comportamiento de un ciclo, tendría sentido eliminar la tendencia, desestacionalizar la serie y suavizarla, para dejar sólo el ciclo.

En cambio, si el propósito fuera apreciar la tendencia (con la finalidad de hacer proyecciones), tal vez conviniera desestacionalizar la serie primero y eventualmente suavizarla, para luego tratar de hallar la función de regresión más adecuada.

En el ejemplo del estudio de Frenkel y S. Rosada mencionado en la primera parte, en cambio, habría que eliminar tendencia y estacionalidad –y si fuese posible el ciclo– para aislar y apreciar la influencia de un hecho puntual, tal como lo fue la reducción arancelaria.

Buenos Aires, DIC/2002

por **Horacio Chitarroni**

Investigador Principal, Área Empleo y Población, IDICSO, USAL.

Email: hchitarroni@siempro.gov.ar